Data, Statistics, and Probability Grade by Grade

Kindergarten

At the kindergarten level, the Grade Level Expectations specify two stems that address (1) interpreting representations, and (2) analyzing patterns, trends, and distributions in data.

Interpreting Representations

The first stem relates to interpreting representations. In kindergarten, students' work with Data, Statistics, and Probability is primarily concerned with interpreting data presented in models and tally charts created by the class.

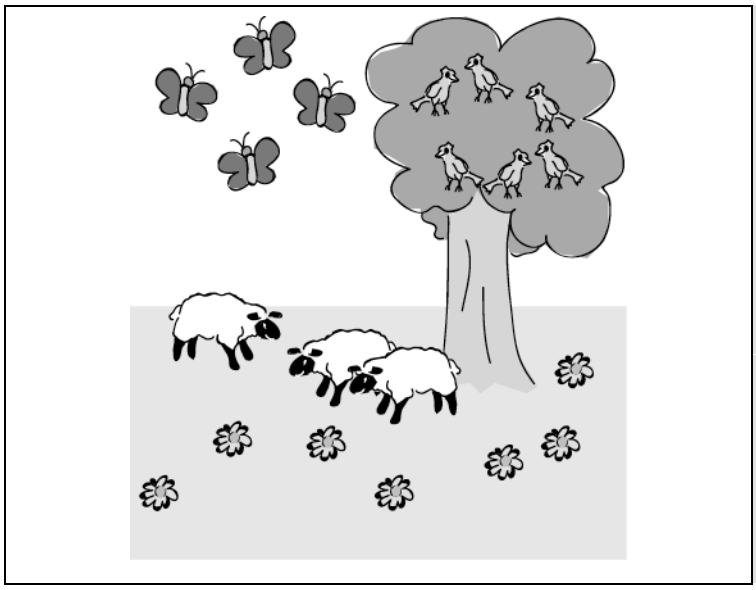
Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. Students need the opportunity to tackle simple problems about data and these problems need to be drawn from a variety of contexts. Such problems provide students with the opportunity to collect and organize data that is familiar to them. In kindergarten, data analysis will be confined to observing whether the number of case is "more than", "fewer than", or "equal to". Here is an example:

Counting Sheep

(a) Write the number of each in the table, below:

(b)



Object:	Number:
Lamb	
Butterfly	
Bird	
1 Flower	

Counting Sheep Chart

(b) Make a graph to show these numbers:

✓			
✓			
✓			
Lamb 	**Butterfly** 	**Bird** 	**1 Flower** 

In the picture, are there more butterflies than flowers?

In the picture, is the number of birds equal to the number of lambs?

Grade 1

At the grade 1 level, the Grade Level Expectations specify three stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) and probability.

Interpreting Representations

The first stem relates to interpreting representations. In grade 1, students' work with Data, Statistics, and Probability focuses on interpreting data presented in representations (models, tally charts, pictographs with one-to-one correspondence, and tables) or collected by the class. Students in grade 1 should know how to represent data in multiple ways. For example, they need the opportunity to represent the same data in tables and in graphs. Students need the opportunity to tackle simple problems about data and these problems need to be drawn from the world about them. Such problems provide students the opportunity to collect and organize data that is familiar to them. When making graphs, students should be encouraged to represent the data both in horizontal and vertical form so that students do not learn that a graph needs to be drawn in one particular way.

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 1, data analysis focuses exclusively on students' observing whether the number of cases is "more than", "fewer than", or "equal to".

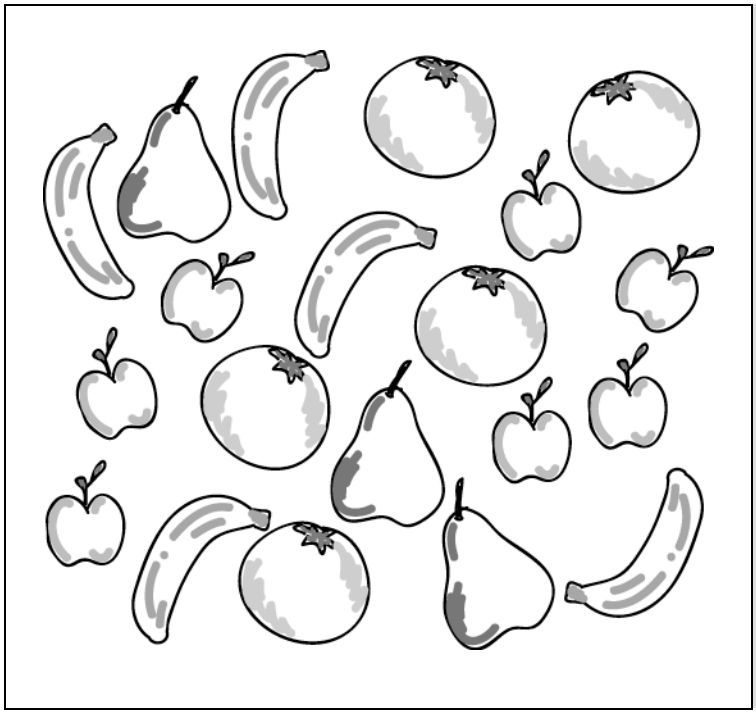
Probability

The third stem relates to probability. Grade 1 students' work in probability is focused on using experiments so that students can describe the likelihood or chance that events are "more likely," "less likely", or "equally likely".

The following two examples show the kind of activities that students could be asked to do in grade 1 to show what they know and can do with Data and Statistics:

Fruits

Count the fruits in the picture and write the number in the table below:



Object:	Number:
	
	
	
	

Fruits Chart

(b) Now make a graph to show the numbers of fruits:

✓			
✓			
✓			
Apple 	**Banana** 	**Orange** 	**Pear** 

Are there fewer apples than banana in the picture?

Is the number of oranges equal to the number of bananas in the picture?

Grade 2

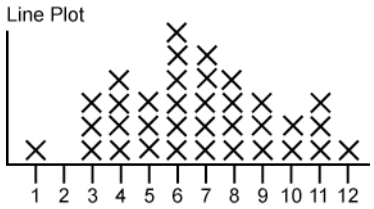
Data, Statistics, and Probability

At the grade 2 level the Grade Level Expectations specify five stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) using counting techniques to solve contextualized problems, (4) probability, and (5) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 2, students' work with Data, Statistics, and Probability focuses on interpreting data presented in representations (pictographs with one-to-one correspondence, line plots, tally charts, or tables) or collected by the class. Students in grade 2 should know how to represent data in multiple ways. For example, they need the opportunity to represent the same data in tables and in graphs. Students need the opportunity to answer questions about the data, and these problems need to be drawn from the world about them. Such problems provide students the opportunity to collect and organize data that is familiar to them. When making graphs, students should be encouraged to represent the data both in horizontal and vertical form.

It is important to note that, in grade 2, the Grade Level Expectations introduce the *line plot*, which is a graph that displays marks (e.g., X or a dot) above a number on a number line to show the frequency of data. Here is an example of a line plot:



Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 2, data analysis focuses exclusively on students observing whether the number of cases is "more than", "fewer than", or "equal to".

Counting Techniques to Solve Contextualized Problems

The third stem relates to counting techniques to solve contextualized problems involving combinations. Students use a variety of counting techniques (e.g., student diagrams, organized lists, tables, or tree diagrams) to solve problems. For example: How many ways can you make 50 cents using nickels, dimes, and quarters?

Probability

The fourth stem relates to probability. In grade 2, students' work in probability is focused on using experiments so that students can describe the likelihood or chance that events are "more likely", "less likely", "equally likely", "certain", or "impossible".

Experimental Design

The fifth stem relates to experimental design. By the end of grade 2, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student-generated question and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the original question and draw conclusions from the data and make predictions.

The following example shows what students could be asked to do in grade 2:

On the Farm

Complete the table below, using both tally marks and numbers:

Object:	Tally:	Number:
cows		
flowers		
trees		
houses		
clouds		

In the picture, are there fewer houses than flowers?

In the picture, is the number of trees equal to the number of clouds?

On the Farm Chart

Now show these numbers in the chart below: (the first one has been done for you).

✓				
✓				
Cows 	**Flowers** 	**Trees** 	**Houses** 	**Clouds** 

Grade 3

At the grade 3 level, the Grade Level Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) using counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 3, students need to be given the opportunity to create and interpret line plots, tally charts, tables or bar graphs. Students need the opportunity to use these graphs and charts to answer questions, formulate conclusions and make predictions related to the data.

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 3, students are expected to analyze data using the terms “most frequent”, “least frequent”, “largest”, and “smallest”. Students need experience in analyzing data with these terms in order to prepare for work in later grades where they will be expected to find and use, for example, measures such as mean, median, mode and range.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In grade 3, students will be expected to identify or describe representations that best display a set of data. Students must organize and display data using tables, tally charts, and bar graphs. They must have opportunities to answer questions related to the data, to analyze the data to formulate conclusions, to make predictions, and to solve problems.

Counting Techniques to Solve Contextualized Problems

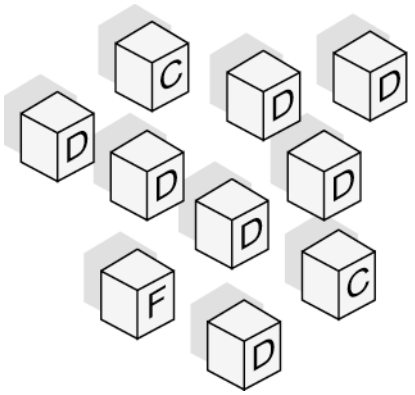
The fourth stem relates to counting techniques to solve contextualized problems. In grade 3 students are expected to use counting techniques such as organized lists, tables, or tree diagrams to solve problems involving combinations and simple permutations.

Probability

The fifth stem relates to probability. Students' work with probability in grade 3 is focused on students use of experiments to determine whether the likelihood of the occurrence of an event is "more likely", "less likely", "equally likely", "certain", or "impossible". The following example shows a grade appropriate task:

Cubes

Rebecca put 10 cubes in a paper sack. 7 of the cubes had the letter 'D' written on them, 2 of the cubes had the letter 'C', and 1 cube had the letter 'F'.



Rebecca closed her eyes and took one cube out of the sack. Is it certain, likely, unlikely, or impossible that the cube she picked had the letter 'D' written on it?

Students are expected to use terms such as "more likely", "less likely", "equally likely", "certain", or "impossible" to predict the likelihood of the occurrence of an event and then use simple experiments to test and investigate their predictions. In grade 3 students can engage with the ideas of probability by working to determine whether or not a simple game is fair.

Experimental Design

The sixth stem relates to experimental design. Throughout grade 3, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student-generated question and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the question. They must also have opportunities to draw conclusions from the data and to make predictions.

Grade 4

At the grade 4 level, the Grade Level Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 4, students need to be given the opportunity to create and interpret line plots, tally charts, tables, bar graphs, pictographs, and circle graphs. Students need the opportunity to answer questions, formulate conclusions and make predictions related to the data. In grade 4, students should learn to justify conclusions and solve problems using these data representations.

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 4, students will be expected to use measures of center (mode and median) and the range of the data to analyze patterns and trends. The table below summarizes the definitions and uses of these measures

Measure	Definition and calculation	Example
Mode	The mode is the most frequently occurring value in a data set. The easiest way to identify the mode is to put the values in order, and then identify the value that occurs most often.	The mode of the data set {1,2,2,3,4,5,5,6,6,7,7,7,8} is 7, as it occurs three times.
Median	The median of a data set is the middle number. (Note: If the data set has an even number of values, find the two middle numbers and calculate the average of these two numbers to find the median.)	The median of the data set {2,3,5,6,9) is 5, as it is the middle number on the data set. (Note: the values must be in order.)
Range	The range of the data set is the difference between the largest and the smallest value.	The range of the data set {2,5,6,8,11} is 11 – 2 = 9.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In grade 4, students will be expected to organize and describe data using tables, line plots, pictographs, or bar graphs that best display a set of data. Students will be expected to use these Representations to answer questions related to the data, to analyze the data to formulate or justify conclusions, to make predictions, or to solve problems.

It is important to note here that while students should be introduced to line plots and pictographs in earlier grades, it is expected that in grade 4 they will be able to create these types of representations for a given data set, as well as to analyze the data and draw conclusions.

Counting Techniques to Solve Contextualized Problems

The fourth stem relates to counting techniques to solve contextualized problems. In grade 4, students are expected to use counting techniques such as organized lists, tables, or tree diagrams to solve problems involving combinations and simple permutations.

The following task shows an appropriate counting problem for students in grade 4:

> **Movie Theater Seats:**
>
> Three friends, Anna, Tasha, and Mae, are going to the movies. They find three seats next to each other in a row in the theater. In how many different ways can they sit?

Probability

The fifth stem relates to probability. In grade 4, students must extend their work in probability by expressing the probability of an event as the ratio of the number of occurrences of an event and the total number of occurrences in the sample space, or probable outcomes. In grade 4, students must predict probability and then test it with an experiment to determine the fairness of a situation or game.

Experimental Design

The sixth stem relates to experimental design. Throughout grade 4, students should be given the opportunity to decide upon the most effective way to investigate a teacher or student generated question and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the question. They must also have opportunities to draw conclusions from the data and to make predictions. In grade 4, students will be expected to ask new questions and connect their analysis with real world situations.

The following example shows a task that students could be asked to solve in grade 4:

Mystery Graphs

Look at the five line plots below: Each graph shows something about a classroom of fourth graders. Which of the five line plots do you think shows:

(a) The number of cavities that the fourth graders have?

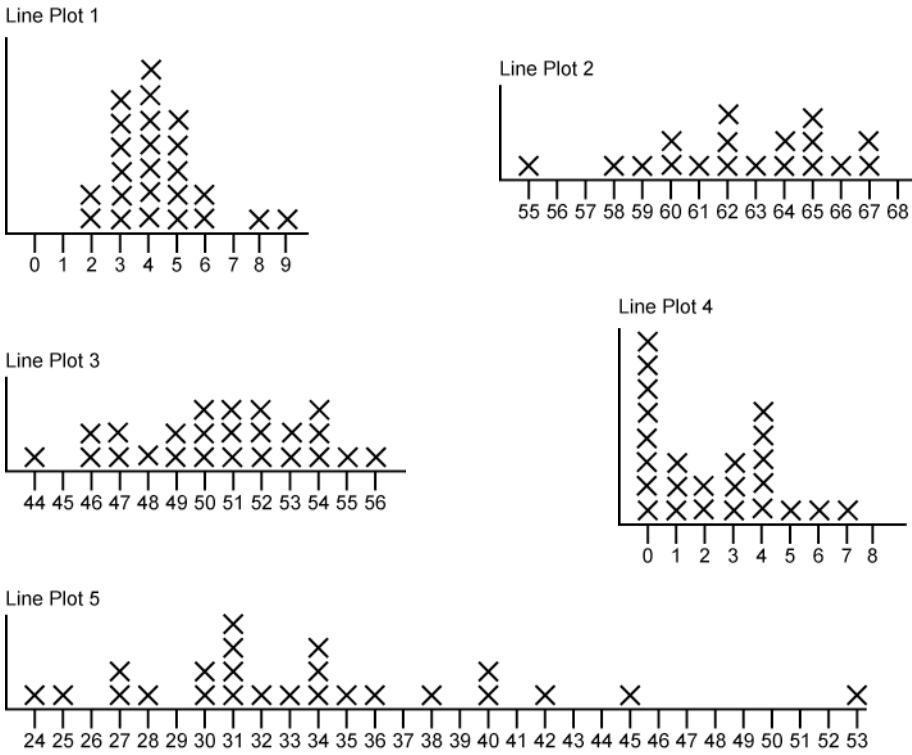
(b) The ages of the fourth graders' mothers?

(c) The heights of the fourth graders, in inches?

(d) The number of people in the fourth graders' families?

(e) Explain why you think the line plot that you picked for c is the one that shows the heights of fourth graders.

(f) Explain why each of the other 4 line plots does not show the fourth graders' heights.



Source: Adapted from Measuring Up

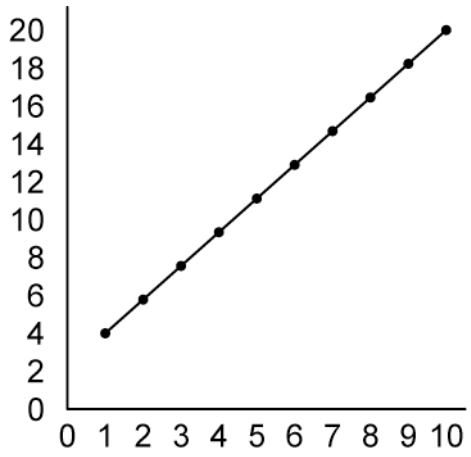
Grade 5

At the grade 5 level, the Grade Level Expectations specify five stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data (4) probability, and (5) experimental design.

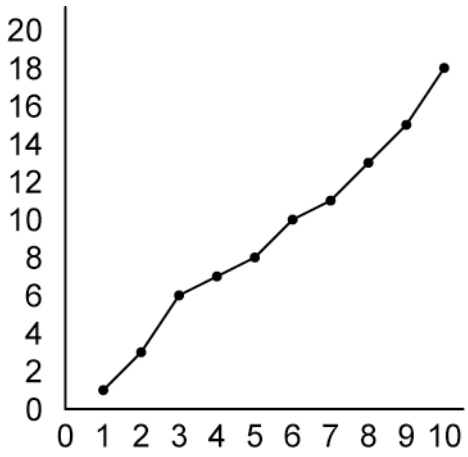
Interpreting Representations

The first stem relates to interpreting representations. In grade 5, students need to be given the opportunity to create and interpret tally charts, tables, bar graphs, circle graphs, and line graphs. Students need the opportunity to answer questions, formulate conclusions and make predictions related to the data. In grade 5, students should learn to justify conclusions and solve problems.

It is important to note that the Grade Level Expectations introduce the *line graph* in grade 5. While in earlier grades students are expected to be familiar with *line plots*, now they are expected to interpret and analyze data displayed in a line *graph*, which is a graphical representation using points connected by line segments to show how some quantity changes over time. This is the beginning of students' work with two-dimensional data sets, which will continue through the later grades. The following graphs are examples of a line graph. Note that in Graph A the points all lie along the same line and in Graph B the points do not lie along the same line. In Graph B, because the line segments joining the points do not lie along the same line the graph itself is not a line.



Graph A



Graph B

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 5, students extend their understanding of measures of center by working with the *mean*, or average of a data set, as well as the median and mode. Students are expected to use measures of center and the range of a data set to analyze situations and solve problems. The table below summarizes these measures:

Measure	Definition and calculation	Example
Mode	The mode is the most frequently occurring value in a data set. The easiest way to identify the mode is to put the values in order, and then identify the value occurring the most often.	The mode of the data set {1,2,2,3,4,5,5,6,6,7,7,7,8} is 7, as it occurs three times.
Median	The median of a data set is the middle number. (Note: there is no middle number if the data set has an even number of values. Calculating the median of an even set of numbers requires finding an average, which is introduced in grade 5.)	The median of the data set {2,3,5,6,9) is 5, as it is the middle number on the data set. (Note: the values must be in order.)
Mean	The mean or average of a set of data is the sum of the values of a set divided by the number of values in the set.	The mean of the data set {1,3,5,8,10} is $\frac{1+3+5+8+10}{4} = \frac{27}{4} = 6.75$
Range	The range of the data set is the difference between the largest and the smallest value.	The range of the data set {2,5,6,8,11} is 11 – 2 = 9.

The following task is an example of a task that students could be asked to solve in grade 5:

Siblings

A class of 25 students was surveyed about how many siblings they each have. The data collected is shown below:

1, 2, 1, 0, 3, 2, 1, 0, 1, 1, 4, 2, 2, 0, 1, 4, 3, 1, 2, 1, 3, 5, 0, 1, 1

Use this set of data to answer the following questions:

(a) What is the mode of the data set? What does this tell you about students and their siblings?

(b) What is the average number of siblings? Is the average the same as the mode?

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In grade 5, students will be expected to identify tables, bar graphs, or line graphs that best display a set of data. Students will be expected to use these displays to answer questions related to the data, to analyze the data to formulate or justify conclusions, to make predictions, or to solve problems.

Probability

The fourth stem relates to probability. In grade 5, students are expected to determine experimental or theoretical probability of events. While in grade 4 they are expected to express probabilities as part whole relationships (for example, 2 out of 5), in grade 5 they are expected to express probabilities as *fractions*. Students must have opportunities to make predictions of the likelihood of events and to test their predictions through experiments. Students should also solve problems in which they determine whether or not a situation or game is fair.

It is important to note that the idea of experimental probability is new in grade 4. While in previous grades students are asked to determine theoretical probabilities, in grade 5 they must determine both experimental and theoretical probability of an event.

Experimental Design

The fifth stem relates to experimental design. Throughout grade 5, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student- generated question and then carry out the investigation. Students should be given the opportunity to collect, organize,

display, and analyze the data in order to answer the question. As well as to draw conclusions from the data and make predictions. In grade 5, students will be expected to ask new questions and connect their analysis with real world situations.

Grade 6

At the grade 6 level, the Grade Level Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 6, students need to be given the opportunity to create and interpret circle graphs, line graphs, or stem-and-leaf graphs.

In grade 6 the stem-and-leaf plot is introduced. A stem-and-leaf plot allows students to organize and display data in order to show its shape and distribution. The plot organizes the data values in terms of “stems” and “leaves”. The “leaf” is usually the last digit, and the “stem” is made up of the digits to the left of the leaf. For example, in the number 41, the digit, 1, is the leaf, and digit 4 is the stem. Below is an example of a data set and it represented in a stem-and-leaf plot:

Data set	Plot	
21, 28, 34, 36, 37, 39, 42, 43, 57, 59	2	1 8
	3	6 7 9
	4	2 3
	5	7 9

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 6 students need to be given the opportunity to use mean, median, mode, and range to analyze situations or solve problems. While students should be introduced to the concept of the range of the data in grade 4, in grade 6 the term *dispersion* is introduced.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In grade 6, students will be expected to organize and display data using tables, line graphs, or stem-and-leaf plots.

As noted above, students in grade 6 should be introduced to the stem-and-leaf plot as a method of representing data. It is important to note that students are also expected to be able to create stem-and-leaf plots, in order to analyze data.

Counting Techniques to Solve Contextualized Problems

The fourth stem relates to counting techniques to solve contextualized problems. In grade 6, students are expected to use counting techniques such as organized lists, tables, tree diagrams, models, and the fundamental counting principle to solve problems involving combinations and simple permutations.

The fundamental counting principle is a rule for determining how many ways one can accomplish two tasks. If there are *m* ways to complete one task, and *n* ways to complete a second task, then there are $m \bullet n$ ways to complete both tasks. For example, if you have 3 kinds of ice cream and 4 toppings, there are 12 ways you can make a sundae.

Probability

The fifth stem relates to probability. In grade 6, students are expected to determine experimental or theoretical probabilities of events in problem-solving situations, and to express probabilities as fractions. Students must be given opportunities to predict theoretical probabilities of events and to test their predictions through experiments and simulations. Students in grade 6 are expected to build on their experiences (from previous grades) in determining whether or not a situation or game is fair. In grade 6 students are also expected to be able to *design* a far game.

Experimental Design

The sixth stem relates to experimental design. Throughout grade 6, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student-generated question or hypothesis and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the question. When appropriate, students should also have opportunities to make predictions, ask new questions, and connect the analysis to real world situations.

The following example shows a task that students could be asked to solve in grade 6:

YogurtTown

During the summer, YogurtTown has a special offer for students: choose one flavor of frozen yogurt, and one topping, and pay just $2.

(a) If there are 5 flavors of yogurt, and 4 different toppings, how many different options are there for the special? Show how you can find your answer using two different methods or models.

(b) If there are 10 flavors of yogurt, and 6 different toppings, how many different options are there for the special?

(c) One day, the store manager began to wonder if students (ages 5 – 18) were visiting the store for his special. The manager asked his customers to tell him their ages. The table below shows the data he collected. Use this data to construct a stem-and-leaf plot.

The table shows the ages of people who visited the store:

10	5	30	12	12	12	13	14	40	42
25	26	15	15	30	30	30	31	40	50
75	76	68	20	21	30	5	3	16	16
15	16	20	19	6	40	6	10	30	6
35	34	12	13	13	70	72	80	80	55
54	16	20	18	18	17	30	5	7	17
15	15	35	5	5	16	16	20	19	10
34	17	18	34	32	18	18	15	15	14

Grade 7

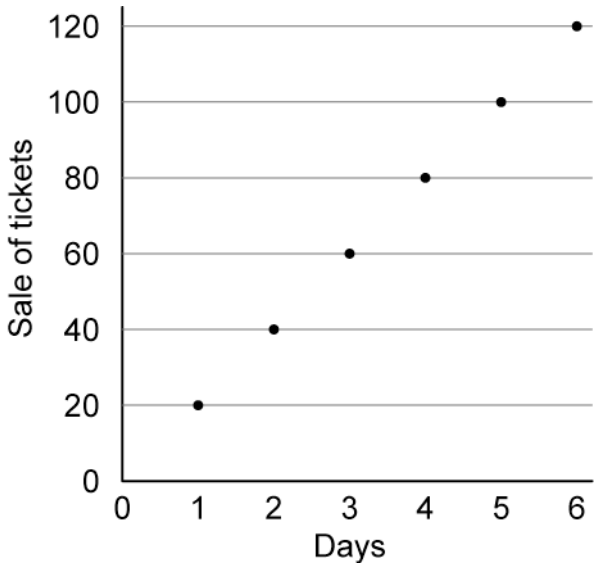
At the grade 7 level, the Grade Level Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 7, students need to be given the opportunity to create and interpret circle graphs, scatter plots, and histograms. Students need to be given the opportunity to analyze data, formulate and justify conclusions, make predictions, and solve problems. In this grade, students are expected to use tables, line graphs, scatter plots, circle graphs to organize and display data.

To extend their knowledge of representations from earlier grades, students in grade 7 should be introduced to scatter plots that represent discrete, linear relationships, and they should be introduced to histograms. In general, a scatter plot refers to a graph that displays paired data by using horizontal and vertical axes. In grade 7, students should have opportunities to interpret and analyze scatter plots that show linear relationships between two data sets. The following graph shows an appropriate graph for grade 7 students:

In this example, as the graph is linear, the number of tickets sold increases the same amount each day. Students working with such graphs should be encouraged to consider when realistic situations should be linear, that is, have a constant rate of change.



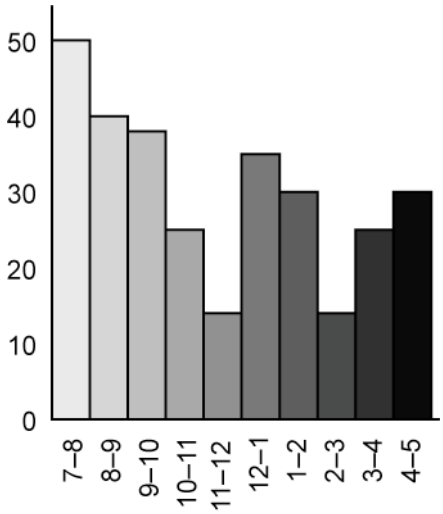
Students should also be introduced to the *histogram*. A histogram is similar to a bar graph, but it specifically refers to a graph that displays the frequency of data values that are found on *continuous intervals*. For example, the *x*-axis of a histogram might display time intervals in which events occurred. An example is shown below:

In a Store

The histogram to the right shows the number of customers who made purchases in a store each hour that the store was open on a given day.

a) In what hour did the most people make purchases?

b) How many people made purchases in this store on that day?



Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In grade 7, students need to be given the opportunity to use the mean, median, mode, and range to analyze situations or solve problems. Students will be expected to determine the effect of *outliers* on mean, median, and mode. By grade 7, students should have had ample experience finding the mean, median, and mode of a data set, and their knowledge of how to calculate these measures should support their ability to analyze the effect of outliers.

An *outlier* is a data value that lies "far away" from the other data values. In statistics, a data point is considered to be an outlier if it lies a certain amount below the first quartile Q1 (the median of the first half of the data) or a certain amount above the third quartile Q3 (the median of the second half of the data). The on-line free encyclopedia, Wikipedia, (http://en.wikipedia.org/wiki/Outlier) defines two kinds of outliers a "mild" outlier and an "extreme" outlier as follows:

Mild outliers

Defining Q_1 and Q_3 to be the first and third quartiles, and IRQ to be the interquartile range ($Q_3 - Q_1$), one possible definition of being far away is:

$$< Q_3 - 1.5 \bullet IQR,$$

or

$$> Q_3 + 1.5 \bullet IQR,$$

Therefore, Q_1 and Q_3 define the so-called inner fences, beyond which an observation would be labeled a mild outlier.

Extreme outliers

Extreme outliers are observations that are beyond the outer fences:

$$< Q_3 - 3 \bullet IQR,$$

or

$$> Q_3 + 3 \bullet IQR,$$

Note that at this grade level, students are not expected to know this formal definition of outlier, as the concept of quartile is not introduced until grade 8. It is, however, useful for teachers to keep this definition in mind as they ask students to consider an observation that is "far away" from the rest of the data.

Consider the following example:

The Effect of an Outlier:

Students in a class took a test, and the following data set represents their scores: {50, 60, 70, 75, 80, 80, 85, 85, 90, 90, 95, 95, 95, 100, 100}. One student was absent, and will take the test tomorrow. If the student scores 100%, how will her score change the mean, or average test score? If the student scores 20%, how will her score change the mean? How low can her score be without changing the average by more than 1 point?

Students are also expected to be able to evaluate the sample from which the statistics were developed. In doing so, students would be expected to consider the ways in which the sample might be biased.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In grade 7, students will be expected to organize and display data using tables, line graphs, scatter plots, circle graphs or stem-and-leaf plots and to use these to answer questions related to the data, to analyze the data to formulate or justify conclusions, to make predictions, or to solve problems. Students are expected to identify or describe the representation or elements of a representation that best display data or a situation. It should be noted that before grade 8, students should have experiences with all of the types of representations listed, except for scatter plots and circle graphs, which are introduced in grade 8.

Counting Techniques to Solve Contextualized Problems

The fourth stem relates to counting techniques to solve contextualized problems. In grade 7, students are expected to use counting techniques such as organized lists, tables, tree diagrams, models, and the Fundamental Counting Principle to solve problems involving combinations and permutations. For example, if there are 15 students in a class, in how many ways can the students line up?

Probability

The fifth stem relates to probability. In grade 7, students are expected to determine experimental or theoretical probabilities of events in problem-solving situations. They must predict the theoretical probability of an event and test their predictions through experiments and simulations. Students are expected to extend their understanding of probability by comparing and contrasting theoretical and experimental probabilities.

Experimental Design

The sixth stem relates to experimental design. Throughout grade 7, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student-generated question or hypothesis and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the question. When appropriate, students should also have opportunities to make predictions, ask new questions, and connect the analysis to real world situations. In grade 7, students are expected to extend their understanding in this area by considering the limitations that could affect their interpretations of the data.

The following task is an example of an experimental design problem that students in grade 7 could be asked to solve:

Conducting an Experiment

Your math teacher is giving a test today, and she is worried that students will not do well on it. When she graded their homework, she noticed that many students had low homework grades.

(a) What kind of experiment could you design to determine if the teacher's prediction is valid? What data will you need?

(b) The data in the table below was taken from another teacher's class. Use this data to determine the mean test score and the mean homework score. Does this support your math teacher's prediction? Why or why not?

(c) Use the data in the table to construct a line graph, using the homework scores as the horizontal axis, and the test scores as the vertical axis. Use your graph to construct an argument in agreement or disagreement with your teacher's prediction.

HW score	0	0	10	50	25	30	50	50	60	75	75	50	55
Test score	25	30	40	50	50	60	60	62	65	70	72	75	75

Table continued:

HW score	80	80	60	85	90	90	100	90	100	100	100	100
Test score	78	80	80	84	86	88	90	92	95	98	100	100

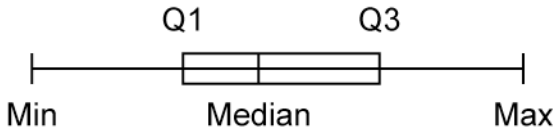
Grade 8

At the grade 8 level, the Grade Level Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. In grade 8, students need to be given opportunities to create and interpret line graphs, scatter plots, histograms, and box-and-whisker plots. Students also need opportunities to analyze data, formulate and justify conclusions, make predictions, and solve problems.

While in grade 7 students are expected to work with scatter plots that show linear trends, in grade 8, they are expected to extend what they have previously learned and to consider scatter plots more generally. They must also be introduced to the *box-and-whisker plot*, which displays the spread of the data by showing the minimum, the maximum, the first quartile (Q1), the median, and the third quartile (Q3) of the data set. The minimum and maximum are easiest to identify if the data set has been ordered. Ordering also helps students identify the median, or the middle number of the data set, as well as Q1 (the median of the first half of the data) and Q3 (the median of the second half of the data). Once these 5 values have been identified, the box-and-whisker plot is created by placing these values on a number line, as shown in the graphic below:



Note that the Inter-quartile range (IQR) is displayed in a box. The minimum and maximum values of the data set make up the “whiskers”. Students can use this kind of data display to answer questions about the dispersion of a data set, or to compare data sets in terms of measures of center and dispersion.

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. In this grade, students will be expected to analyze patterns, trends, or distributions in data in a variety of contexts by using the mean, median, mode, range, quartile values, or estimated line of best fit to solve problems.

In grade 8, the concept of a *quartile* is introduced. The first quartile (Q1) of a data set is the median of the first half of the (ordered) data. The third quartile,(Q3) is the median of the second half of the (ordered data). Students must know how to calculate the quartiles, and how to include them in a box-and-whisker plot, as shown above.

The following task addresses this type mathematics:

Timely Commute

The data below represent the time of Michael's evening commute for 2 weeks in March.

	Time in Minutes
Week 1	35, 43, 32, 36, 40
Week 2	45, 35, 34, 33, 36

(a) Find the median of this data.

(b) Find the upper and lower quartile of this data.

(c) Build a box and whisker plot to represent this data.

In grade 8, students should also have opportunities to extend their knowledge of data representations by learning how to estimate lines of best fit. Given paired data, students should be able to create a scatter plot of the data, and to estimate a line of best fit in order to analyze any trends in the data, or to predict (extrapolate) values not represented in the data set. Students' work in grade 8 (and in grades 9-10) with estimating lines of best fit prepares them to find regression lines in grades 11-12.

In grade 8, students will also be expected to analyze bias in sampling. The following task is an example of a question that students should be asked:

Voters

Denise wants to find out whether or not local voters are in favor of funding a new baseball field. She plans to post surveys around the town with instructions to fill the survey out and mail it to her. Explain why her sampling method may or may not be a good representation of the voting population.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In this grade student are expected to use scatter plots to organize and display data. Students will be expected to use scatter plots and other displays to analyze the data to formulate or justify conclusions, to make predictions, or to solve problems. They must also be able to identify representations or elements of representations that best display a given set of data or situation,

Counting Techniques to Solve Contextualized Problems

The fourth stem relates to counting techniques to solve contextualized problems. In grade 8, students are expected to use counting techniques such as organized lists, tables, tree diagrams, models, and the Fundamental Counting Principle to solve problems involving combinations and permutations. For example, if there are 15 students in a class, in how many ways can the students line up?

Probability

The fifth stem relates to probability. In grade 8, students are expected to determine experimental or theoretical probabilities of events in problem-solving situations. They must also predict the theoretical probability of an event and test their predictions through experiments and simulations. Students are expected to compare and contrast theoretical and experimental probabilities.

Experimental Design

The sixth stem relates to experimental design. Throughout grade 8, students should be given the opportunity to decide upon the most effective way to investigate a teacher- or student-generated question or hypothesis and then carry out the investigation. Students should be given the opportunity to collect, organize, display, and analyze the data in order to answer the question. When appropriate, students should also have opportunities to make predictions, ask new questions, and connect the analysis to real world situations. In grade 7, students are expected to consider the limitations that could affect their interpretations of the data. This continues in grade 8.

Grade span 9-10

At the 9-10 grade span, the Grade Span Expectations specify six stems that address (1) interpreting representations, (2) analyzing patterns, trends, and distributions in data, (3) organizing and displaying data, (4) counting techniques to solve contextualized problems, (5) probability, and (6) experimental design.

Interpreting Representations

The first stem relates to interpreting representations. Students are expected to be able to interpret a given representation including box-and-whiskers plots, scatter plots, bar graphs, line graphs, circle graphs, histograms and frequency charts. Students are expected to be able to use the data to make observations, to answer questions, and to formulate or justify conclusions. Additionally, students are expected to be able to critique conclusions and make predictions. Finally, students are expected to be able to use graphs and charts to solve problems within mathematics or across disciplines or contexts (e.g., media, workplace, social and environmental situations).

Analyzing Patterns, Trends, and Distributions in Data

The second stem relates to analyzing patterns, trends, and distributions in data. Students are expected to be able to analyze patterns, trends, or distributions in data in a variety of contexts. Students are expected to be able to determine, use, and analyze the mean, median, mode, range, outliers, quartile values, or variation in solving problems. Students are also expected to interpret statistical analysis in the light of the sampling techniques deployed.

Organizing and Displaying Data

The third stem relates to organizing and displaying data. In addition to interpreting data in a representation, in grades 9-10, students are expected to be able to organize and display one- and two-variable data. Students are expected to have familiarity with box-and-whisker plots, scatter plots, bar graphs, line graphs, circle graphs, histograms, and frequency charts. Given a set of data, students should be able to identify the best display. For example, students should recognize that continuous data is better represented on line graphs, scatter plots, and box-and-whiskers plots and that categorical data is better represented in bar graphs, circle graphs, histograms, and frequency charts.

The following task shows an example of a problem that students could be asked to do in grades 9-10.

Daylight

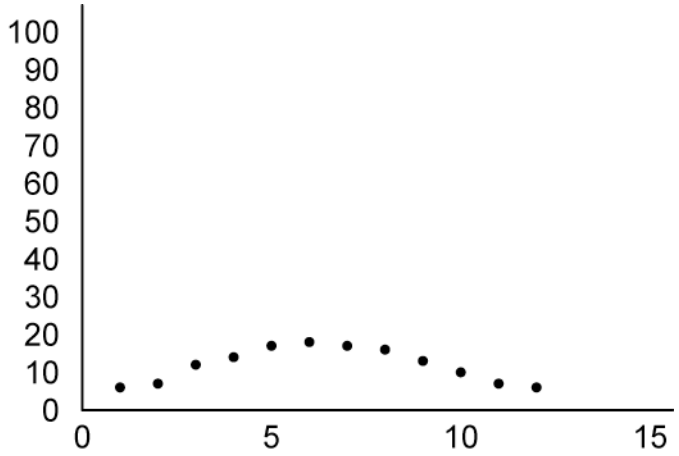
The following table shows the average number of hours of daylight each month during a year in Anchorage, Alaska:

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Hours	6.2	7	12	14.1	17	18.6	17.5	15.7	13	10	7.0	5.8

Construct a graph that displays this data, and justify your decision to choose a particular type of graph.

What does your graph tell you about the data that the table given might not?

A student decided to graph this data using a scatter plot. The student's graph is shown below:



From this graph the student concluded that the number of hours of daylight each day does not fluctuate too much over the course of the year. Critique this student's claim in terms of (1) the graph he drew and (2) how he interpreted the graph.

Counting Techniques to Solve Contextualized Problems

The fourth stem relates to counting techniques to solve contextualized problems. In grades 9-10, students are expected to use counting techniques to solve contextualized problems involving combinations or permutations. Among the counting techniques that students need to be able to utilize are: organized lists, tables, tree diagrams, models and the Fundamental Counting Theorem.

Probability

The fifth stem relates to probability. In grades 9-10, students are expected to solve problems involving probability. Students should know that the probability of an event happening is always between 0 (event is impossible) and 1 (event must always happen). Students should be able to identify the sample space, or the collection of all possible outcomes for an experiment. Here is an example:

Six-Sided Random Number Cube

Consider the rolling of two fair 6-sided random number cubes (RNCs) (with sides labeled 1, 2, 3, 4, 5, and 6). Thirty-six (36) equally likely outcomes are possible. These outcomes are called the sample space.

(a) What is the probability that the sum of the two numbers rolled on the two RNCs is 10?

(b) What is the probability that both cubes land on the same number?

Additionally, students are expected to be well versed in the differences between theoretical probability and relative frequency (often called experimental probability).

Theoretical Probability

The theoretical probability that the sum of the two random number cubes rolled is 10 in the example is 4/36 or 1/9. When the random number cubes are tossed 100 times, what is the probability that the sum of the two random number cubes rolled is 10?

Students should understand how they might go about conducting experimental trials. Additionally, students are expected to understand that the experimental probability may not equal the theoretical probability. However, as the number of trials is increased the experimental probability should increasingly approach the theoretical probability. For example, if a coin is tossed repeatedly, over time the distribution of results will approach 50% heads and 50% tails. This is sometimes referred to as “the Law of Large Numbers”.

Experimental Design

The sixth stem relates to experimental design. In grades 9-10, students are expected to design experiments in response to a teacher- or student-generated question or hypothesis. Students should be able to identify an appropriate methodology (e.g. survey, observations, experimentation) to answer the questions as well as understanding the allowances and constraints of different methodologies. Students should utilized appropriate sampling techniques (e.g., simple random samples and stratified random samples) to collect the data necessary to answer the question. Students are expected to be able to carry out their study design and to collect, organize, and appropriately display the data. Further, students should be able to analyze the data to draw conclusions about

the questions or hypotheses guiding the experiment. Students are expected to understand the limitations of their study and how their limitations affect interpretations. One limitation that students should understand is that it might be impossible to generalize their findings. For example, if a study of ninth graders' favorite movies is conducted at one high school, the findings cannot be generalized to ninth grade students in general since trends may vary from high school to high school. Finally, students should be able to utilize collected data to make predictions, ask new questions and make connections.

Grade span 11-12 for all students

At the 11-12 grade span, the Grade Span Expectations specify five stems that address (1) regression, (2) analyzing measures of dispersion in data, (3) finding regression functions, (4) combinations and permutations, and (5) probability.

Regression

The first stem relates to regression. In grades 11-12, students are expected to understand and utilize regression functions (linear, quadratic & exponent). Students should be able to analyze data to make inferences and to formulate, justify and critique conclusions.

For example, for a linear regression function, students need to understand that the linear regression equation is related to a scatter plot of all of the data. The linear regression equation is a straight line with the equation $y = mx + b$ that best fits the set of data points. While students may estimate lines of best fit or find them using technology (See Finding regression lines, below) students should come to understand that the regression line is considered to "best" fit the data because it is the line that minimizes the distances from the data points to the line.

Students are expected to interpret the slope, *m*, of the regression equation in terms of the relationship between two variables (*x* and *y*). For example, when the slope of the line is positive, there is a positive relationship between the *x*-values and y-values: as the values of one variable increase or decrease, so do the values of the other variables. If the slope of the line is negative, then this relationship is reversed. When the slope of the line approaches 0, there is no relationship between the two variables. Students are expected to be able to use these interpretations to make inferences and to formulate, justify, and critique conclusions.

It is important to note that the slope of the line does not determine the *strength* of the relationship between the two variables. The strength of the relationship is measured by *r*, the *correlation coefficient*. The Grade Level Expectations do not require students to study *r*, though it is common for this concept to appear in high school mathematics texts. The correlation coefficient will be discussed further below in terms of finding regression functions.

Measures of Dispersion

The second stem relates to the analysis of measures of dispersion of data. Students are expected to calculate and analyze measures of dispersion. Students should be able to calculate the *standard deviation* of a given set of data. Students should understand that the standard deviation measures variation by indicating how far, on average, data points are from the mean. For example, given the heights of members of a high school basketball team, students should be able to find the mean height, and find the standard deviation, or how far, on average, the actual heights of players are from the mean height. Students are

also expected to understand how to calculate and use measures of variance and percentiles.

The following problem is an example of a task that students in grades 11-12 could be asked:

Standard Deviation

The population standard deviation of a data set of 100 numbers is 4.2. If you calculate the sample standard deviation, *s,* how will your answer compare to the population standard deviation? Justify your answer.

Finding Regression Functions

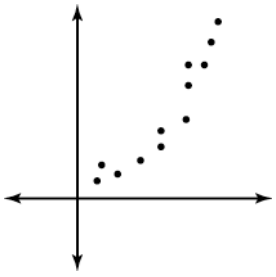
The third stem relates to finding regression functions. In addition to being able to *interpret* regression functions, students are expected to be able to find or estimate linear, quadratic and exponential regression functions. Students should be able to do this by displaying data in a scatter-plot and using their understandings of properties of the regression equation. For example, given a set of data points students should be able to estimate the line of best fit for a linear relationship on the graph and from their estimation, determine both the *y*-intercept and the slope of the line of best fit. From this information, students can estimate the regression function and find the regression equation using technology.

As noted above, the strength of the relationship between two variables is measured by *r*, the correlation coefficient. Calculating this value by hand is cumbersome, but students can use technology (e.g. graphing calculator or spreadsheet) to find *r*, which will always be a value between -1 and 1. If *r* is close to -1, then there is a strong, negative relationship. If r is close to 1, then there is a strong relationship. If *r* is close to 0, then there is only a weak, or no relationship between *x* and *y*. While, again, the Grade Span Expectations do not require students to be familiar with *r*, it is a useful way to evaluate the strength of the relationship.

The following example asks students to connect their understanding of a scatter plot and the correlation coefficient:

Correlation Coefficient

Consider the data graphed to the right.
Which one of the following numbers could be the correlation coefficient for a linear regression for the data graphed to the right?



(a) $r = -.74$	(b) $r = .08$	(c) $r = .41$	(d) $r = .72$

The following problem is an example of a regression task that is appropriate for students in grades 11-12.

Hotdogs

The table on the right shows the calories and sodium content of beef hotdogs of various brands.

(a) Use your graphing calculator to find the linear correlation coefficient, and the equation of the regression line:

(b) If a hotdog had 150 calories, what would the sodium content be, according to your regression line?

(c) If a jumbo hotdog had 300 calories, what would the sodium content be, according to your regression line?

(d) Why is your answer in part (b) more trustworthy than your answer in part (c)?

Calories	Sodium (Mg)
186	495
181	477
176	425
149	322
184	482
190	587
158	370
139	322
175	479
148	375
152	330
111	300
141	386
153	401
190	645
157	440
131	317
149	319
135	298
132	253

Combinations and Permutations

The fourth stem relates to combinations and permutations. Students are expected to be able to solve problems involving combinations and permutations involving a variety of strategies. Students should be familiar with the following concepts:

Concept	Definition and calculation	Example
Factorial: n!	$n! = n \bullet (n-1)(n-2) \bullet \bullet \bullet 3 \bullet 2 \bullet 1$ "n factorial" is the product of n and each integer from n down to 1.	Five students are getting in line for lunch. In how many ways can they stand in line? $5! = 5 \bullet 4 \bullet 3 \bullet 2 \bullet 1 = 120$
Permutation: nPr	$nPr = \frac{n!}{(n-r)!}$ The permutation formula allows you to find the number of ways to arrange r items from a set of n items. Here, order matters.	For example, if we are arranging pairs of letters from the set *a*, *b*, *c*, *d*, then *ab* and *ba* are different permutations.
Combination nCr\\	$nCr = \frac{n!}{r!(n-r)!}$ This formula allows you to find the number of combinations of r items from a set of n items. Order does not matter.	For example, if we are making combinations of pairs of letters from the set *a*, *b*, *c*, *d*, then *ab* and *ba* are counted as the same combination.

Students should be able to solve a problem such as:

Ice Cream

An ice cream shop sells 32 different types of ice cream. Alyssa wants to have 2 scoops of ice cream (each scoop a different flavor). How many combinations of two different types of ice cream exist at this ice cream shop?

Additionally, students are expected to be able to find the union, intersection and complement of sets. Given two sets of numbers, these terms are defined as follows:

Term	Definition	Notation	Example
Union	The union of two sets A and B is the set of all items that are members of set A or set B	$A \cup B$	The union of the sets {1,2,4,5} and (1,2,7,9} is the set {1,2,4,5,7,9}.
Intersection	The intersection of two sets A and B is the set of all items that are members of both set A *and* set B.	$A \cap B$	The intersection of the set of prime numbers and the set of even numbers is {2}.
Complement	Given a set A and a universal set U, the complement of set A is all of the items in the "universal" U that are *not* in set A	$A^{\complement}$	Given the universal set of the Real Numbers, the complement of the Rational numbers is the set of Irrational numbers.

Probability

The fifth stem relates to probability. Students are expected to be able to design and critique experimental models in order to approximate desired probabilities. Thus, students are expected to be able to understand that repeated experimental trials lead to an approximation of the theoretical probability. For example, when 2 fair Random Number Cubes (with sides labeled 1, 2, 3, 4, 5, 6) are rolled, the theoretical probability that the sum of the numbers rolled is 10 is 1/9. Repeated experimental trials, will lead to finding a probability of rolling a sum of 10 that more and more closer approximates the theoretical probability.

Additionally, students are expected to be able to solve probability problems by applying concepts of counting, random variables, independence/dependence of events, and conditional probability. For example, if two events are unrelated then the probability of either event happening is the sum of the probability of each event happening.

The following problems are examples of probability tasks that students should be able to do in grades 11-12.

Almost Fair

Sometimes the probability of winning a game is close to 1/2 but is not exactly equal to 1/2. Such a game can be called *almost fair*.

Suppose the game is this:

You toss one Random Number Cube 4 times.
If you don't get a 6 on any of the tosses, you win.
Show that this game is *almost fair.* What is the probability of winning this game?

Three more examples are provided below:

Random Number Cubes

You toss two Random Number Cubes (RNC) *n* times.
If you don't get a double 6 on any of the tosses, you win.
What should *n* be to make this game *almost fair*?
(In other words, how many tosses of the two RNCs should there be to make this an almost fair game?) What is the probability of winning this game?

Chords in a Circle

In a circle, pick a chord at random. What is the probability that its length will be greater than its radius?

Triangles

A stick is cut at random in two places. What is the probability that the three pieces can form a triangle?

Grade Span 11 –12 for Prospective Mathematics Majors

In addition to the mathematics identified in the grade span 11-12, students who plan to major in mathematics, engineering, or the sciences need to study additional mathematical concepts.

Measures of Dispersion

Students must be able to analyze and interpret measures of dispersion and central tendency for normal distributions. Among the measures of dispersion that students must be able to utilize and interpret are standard deviation, variance, and percentiles.

Here is an example:

Standardized Test

On a certain standardized test, the mean is 115.

If a score of 180 is exactly 2.5 standard deviations above the mean, what is the standard deviation?

What score falls exactly 2.5 standard deviations *below* the mean?

Least Squares Regression

Students must understand the method of least squares for linear regression. Students should develop this understanding through the use of technology which models that the line of best fit is the straight line that has the smallest possible sum of squared errors. Students must be able to use technology and the least squares regression method in order to find the regression equation.